REPORTS OF THE FINNISH ENVIRONMENT INSTITUTE I | 2018

Interlaboratory comparison 11/2017

Bacterial toxicity test using bioluminescent bacteria

Johanna Järvistö, Katarina Björklöf and Markku Ilmakunnas

Finnish Environment Institute

REPORTS OF THE FINNISH ENVIRONMENT INSTITUTE I | 2018

Interlaboratory comparison 11/2017

Bacterial toxicity test using bioluminescent bacteria

Johanna Järvistö, Katarina Björklöf and Markku Ilmakunnas

Helsinki 2018
Finnish Environment Institute

REPORTS OF THE FINNISH ENVIRONMENT INSTITUTE 1 | 2018
Finnish Environment Institute SYKE
Proftest SYKE

Layout: Markku Ilmakunnas

The publication is also available in the Internet: www.syke.fi/publication | helda.helsinki.fi/syke

ISBN 978-952-11-4903-0 (pbk.)
ISBN 978-952-11-4904-7 (PDF)
ISSN 1796-1718 (print)
ISSN 1796-1726 (Online)

Author(s): Johanna Järvistö, Katarina Björklöf and Markku Ilmakunnas

Publisher and financier of publication: Finnish Environment Institute (SYKE)
P.O. Box 140, FI-00251 Helsinki, Finland, Phone +358 295 251 000, syke.fi.

Year of issue: 2018

SYKE

ABSTRACT

Interlaboratory comparison 11/2017

The ecotoxicological laboratory of SYKE and Proftest SYKE carried out this interlaboratory comparison for analysis of ecotoxicity in liquids using *Aliivibrio fischeri* (formerly *Vibrio fischeri*)– bacterial test. One clear synthetic sample and one dark as well as coloured sample were sent in October 2017 (BTOX 11/2017) to eight participants. In total participants provided 22 results.

The mean of individual results of the participants were used as assigned values for the clear sample. As expected, all participants but one reported satisfactory results. No evaluations of performances of the colored sample were done. The biggest problems related to credible results were related to pH adjustment of the samples and the non-toxicity of the sample which required extrapolation of the measured results in order to receive an exact value for the EC50%.

Warm thanks to all the participants of this interlaboratory comparison!

Keywords: *Aliivibrio fisheri*, kinetic luminescent bacteria test, acute toxicity, coloured and turbid samples, interlaboratory comparison , quality control, ecotoxicology

TIIVISTELMÄ

Laboratorioiden välinen vertailumittaus 11/2017

SYKEn ekotoksikologinen laboratorio ja Proftest SYKE järjestivät akuuttia valobakteeritestiä suorittaville laboratorioille vertailumittauksen lokakuussa 2017 (BTOX 11/2017). Vertailumittaukseen osallistui yhteensä kahdeksan osallistujaa 22 testituloksella.

Kirkkaan näytteen mittaussuureen vertailuarvona käytettiin osallistujien yksittäisten tulosten keskiarvoa. Odotetusti kaikkien, paitsi yhden, osallistujan tulokset olivat hyväksyttäviä. Osallistujien pätevyyksiä ei arvioitu värillisen näytteen kohdalla. Suurimmat tuloksien epävarmuuteen vaikuttavat tekijät olivat näytteen pH:n säätö ja näytteen vähätoksisuus, mikä saattoi kasvattaa tulosten vaihtelua tuloslaskennan yhteydessä.

Lämmin kiitos vertailumittauksen osallistujille!

Avainsanat: *Aliivibrio fisheri*, kineettinen valobakteeritesti, akuutti myrkyllisyys, värilliset ja sameat näytteet, vertailumittaus, laadunvarmistus, ekotoksikologia

SAMMANDRAG

Interkalibrering 11/2017

SYKEs ekotoxikologiska laboratorium arrangerade med SYKE Proftest en interkalibrering för akut luminescent *Aliivibrio fisheri* -bakterietest i oktober 2017 (BTOX11/2017) med ett klart syntetiskt prov och ett grumligt prov. Åtta deltagare deltog med 22 resultat.

Som väntat, var alla deltagares resultat utom en, godkända. Det färgade provet var inte tillräckligt homogent och deltagarnas kompetens värderades inte. De största problemen med det färgade provet var justeringen av pH och provets låga toxicitet, som lätt kan orsaka räknefel vid extrapoleringen av resultaten.

Ett varmt tack till alla deltagarna i testet!

Nyckelord: *Aliivibrio fisheri*, akut toxicitetstest. kinetisk test med luminiscerande bakterier, bakterietest, färgade och grumliga prover, kvalitetskontroll, ekotoxikologi

CONTENTS

1 Introduction

Ecotoxicity testing using *Aliivibrio fischeri* (formerly *Vibrio fischeri*) is a simple and cost effective testing method suitable for testing of a variety of waters, sediments and soil samples. The method is based on a bioluminescent enzyme produced in the basic metabolism of the marine *A. fischeri* -bacteria and inhibition of bioluminescence produced when the bacteria are stressed. Ecotoxicity testing using *Aliivibrio fischeri* is widely used and the methods are standardized by the International Standardization Organization. Two test methods, which both use freeze dried bacteria, were used in this interlaboratory comparison; the traditional method [1] and the kinetic method [2].

The standard procedure includes the use of a reference substance, usually 3,5-dichlorophenyl (3,5-DCP), to monitor the viability of the test bacteria used. 3,5-DCP is a clear solution that gives easily repeatable results with a straightforward inhibition response. However, the scope of the kinetic test method also includes samples that are colored or turbid. These properties might interfere with the light detection and distort the test results.

This interlaboratory comparison for analysis of exotoxicity to *A. fischeri* in two samples was arranged in October 2017 (BTOX 11/2017). The Finnish Environment Institute (SYKE) is appointed National Reference Laboratory in the environmental sector in Finland. The duties of the reference laboratory include providing interlaboratory proficiency tests and other comparisons for analytical laboratories and other producers of environmental information. This interlaboratory comparison has been carried out under the scope of the SYKE reference laboratory and it provides an external quality evaluation between laboratory results and mutual comparability of analytical reliability. It was carried out applying, when suitable, the international guidelines ISO/IEC 17043 [3] and ISO 13528 [4]. The Proftest SYKE has been accredited by the Finnish Accreditation Service as a proficiency testing provider (PT01, ISO/IEC 17043, www.finas.fi/sites/en, ISO/IEC 17043). The organizing of this interlaboratory comparison is not included in the accreditation scope.

2 Organizing the interlaboratory comparison

2.1 Responsibilities

Organizer and analytical expert:	Finnish environment Institute (SYKE) Laboratory Centre, Ecotoxicology and Risk Assessment Johanna Järvistö, johanna.jarvisto@ymparisto.fi puh. +358 295 251 243
Interlaboratory comparison coordinator:	Finnish environment Institute (SYKE) Proftest SYKE Katarina Björklöf, katarina.bjorklof@ymparisto.fi puh. +358 40 148596

2.2 Participants

In total eight participants took part in this interlaboratory comparison (Table 1). Participants were mainly from Finland, with one participant from Turkey. The participants include commercial and other laboratories as well as a reagent manufacturer.

The test is possible to perform with either a single tube luminometer or a well plate reader with luminometer features. The test method itself also offers two options for measurement. The standard method relies on a single measurement result as opposed to the kinetic method, which utilizes continuous measurement for several seconds and a maximum value is used for calculations. The procedures of adding the bacteria suspension also differ between the methods. The luminometer automatically injects the bacteria suspension to the sample in the kinetic method while it is pipetted in manually in the standard method.

All equipment and test method combinations were accepted in this interlaboratory comparison. The aim was to compare the equipment and test type in cases where the amount of results was enough for statistical calculations. Altogether three of the participants used accredited analytical methods at least for a part of the measurements.

For this interlaboratory comparison, the organizing laboratory (T003, www.finas.fi/sites/en, ISO/IEC 17025) has the participant code 8. The standard method is included in the scope of accreditation of the organizing laboratory and the kinetic method has been accepted in the scope of accreditation in November 2017. Both methods were performed in a single tube luminometer and the method used was chosen according to the properties and type of any given sample.

Table 1. Participants of the luminescent bacteria interlaboratory comparison BTOX 11/17.

Country	Organization	Number of results
Finland	Aboatox Oy, Turku	2
	Eurofins /NabLabs Oy Kokkolan yksikkö	4
	Eurofins Environmenta Testing Finland Oy, Lahti	4
	Finnish Environment Institute, Jyväskylä	4
	Kokemäenjoen vesistön vesiensuojeluyhdistys	2
	Kymen Ympäristölaboratorio, Kuusankoski	1
	Metropolilab Oy	1
Turkey	Istanbul Water and Sewage Administration, 34060 Eyup-ISTANBUL	4

2.3 Samples and delivery

This test scheme consisted of two samples, a clear sample of 3,5-DCP as well as a colored and turbid sample that was prepared using a malt beverage. The clear sample was included to determine the differences in actual test methods (accuracy) calculation of the results, whereas the colored sample was the sample created to challenge the testing procedures.

The clear sample was prepared on 6th of October 2017 by weighing 20.31 mg of 3,5-DCP and dissolving it in 1000 ml of ultra-pure water. Samples were divided into 100 ml batches in 125 ml PP plastic bottles and stored in +4 °C until delivery.

The colored sample was prepared from a malt beverage on 13th of October 2017 and was aerated in room temperature for 72 hours, divided into 100 ml batches in 125 ml PP plastic bottles and sterilized in 121°C for 15 min. After cooling, the samples were stored in +4°C until delivery.

The results were requested to be reported as a percentage of original samples instead of actual concentration. Therefore no testing of actual concentrations was needed. All test results were requested as EC50-values in percent of original sample both for 15 min and 30 min exposure.

Both samples were labeled and packaged on 16th of October and posted on the same day. The samples arrived to the participants latest on 18th of October 2017. None of the participants reported problems of any kind about sample delivery. The participants were asked to analyze the samples by the 20th of October 2017. The preliminary results were delivered to the participants via email on 17 November 2017.

2.4 Feedback from the interlaboratory comparison

Feedback was gathered throughout the process and most of it was received during sample analysis and result reporting. The feedback that required action will be taken under consideration and we will aim to improve on the next round.

2.5 Processing the data

2.5.1 Homogeneity and stability of the samples

The results of homogeneity and stability testing of the samples indicate that the clear samples were relatively stable (Table 2). However, the homogeneity and stability of the colored samples were not satisfactory (Table 3). The homogeneity results of the colored sample in the interlaboratory comparison were not similar to results obtained during pretesting of the samples. Reasons for this are considered in Chapter 3.2

FEEDBACK FROM THE PARTICIPANTS

Participant	Comments	Action / Proftest
1	The pHs of the samples were acidic. The participant neutralized the pH according to normal procedures. The participant commented that the pretreatment of samples may be a serious source of variation between results if not all participants have neutralization of acidic samples as their normal pretreatment procedures.	Thank you for your comment. This year the aim was to bring more realism to also the sample pretreatment step so the clear sample was prepared differently from before and these changes resulted in a lower pH value than anticipated. The pH issue will be taken into consideration when preparing for the next round. The dark sample was chosen for other properties and was also aimed to be more challenging sample.
1	Inhibition in the strongest dilution (50%) was 34 %. EC50 value could not be determined.	This may be the case if the samples are only slightly toxic. Then EC50 values are usually not reported, but the results are given as participant 1 has reported.
2	The participant reported their results incorrectly. The corrected results are: Sample: Clear EC50(15min)=20.3 % ± 0.8 %, Uncertainty 2s = 1.6 % EC50(30min)=19.0 % ± 1.0 %, Uncertainty 2s = 2.0 % Sample: Color EC50(15min)=0.5 % ± 0.1 %, Uncertainty 2s = 0.3 % EC50(30min)=0.7 % ± 0.1 %, Uncertainty 2s =0.3 %	The policy of interlaboratory comparisons does not allow corrections of the participant's results after the preliminary results are sent. The EC50 for clear sample result was satisfactory if it would have been reported correctly. The participant can re-calculate the z scores according to the Guide for participants [5].
4	Incubation at room temperature (incubator broke down).	This may have affected the result. The result was included in the statistic calculation of the results.
5	The equipment will not report EC50-value, 15 min values.	The EC50-value, 30 min- results are more widely used. EC50-value, 15 min-are usually additional information.
7	The result sheet form in ProftestWEB was in the wrong unit (mg/l).	The units for reporting results were corrected to %. The organizer apologizes for this error.

Table 2. Stability and homogeneity testing of the clear sample.

Date	EC50 (%), 30 min Parallell results			Mean	Sd
	a	b	c		
6.10.2017	17			17	
11.10.2017	18	16		17	1.5
16.10.2017	14	16	18	16	1.7
19.10.2017	18			18	
20.10.2017	16			16	
				17	1.6

Sd: the standard deviation.

Table 3. Stability and homogeneity testing of the colored sample.

Date	EC50 (%), 30 min Parallell results		Mean	Sd
	a	b		
10.10.2017	61		61	
12.10.2017	55	91	73	25
18.10.2017	32	31	32	0.4
20.10.2017	73		73	
22.11.2017	79	51	65	19
			61	15

Sd: the standard deviation.

2.5.2 Pretesting the data

The normality of the data was tested by the Kolmogorov-Smirnov test. The outliers were rejected according to the Grubbs or Hampel test before statistical calculations. More information about the statistical handling of the data is available in the Guide for participants [5].

2.5.3 Assigned value of clear sample

The mean value of the results reported by the participants was used as the assigned value for the clear sample (Table 4). The results of the colored sample varied a lot. Therefore and assigned values were not given.

When the mean value was used as the assigned value, the expanded measurement uncertainty was estimated based on the standard deviation [3, 4]. According to proficiency test criteria, both assigned values were reliable [4].

Table 4. The assigned values and their uncertainties in the luminescent bacteria interlaboratory comparison (BTOX 11/2017).

Measurand	Sample	Unit	Assigned value	U_{pt}	U_{pt}, %	Evaluation method of assigned value
EC50 value, 15 min	Clear	%	17.3	1.8	10	Mean
EC50 value, 30 min	Clear	%	17.8	2.1	12	Mean

U_{pt} = Expanded uncertainty of the assigned value

After reporting the preliminary results no changes have been done for the assigned values.

2.5.4 Standard deviation for proficiency assessment and results' evaluation

The standard deviation for proficiency assessment was set only for the clear sample EC50, 30 min result. For the clear sample EC50 15 min, the data set was too small (n<6) for setting a standard deviation for proficiency assessment. For the colored sample performance evaluations were not done. This was mainly because the homogeneity of the samples was not sufficient.

For the EC50 30 min results, the performance of each participant was expressed as z scores (Appendix 2 and 4).

For EC50 15 min results, where the number of reported results was low (n<6), the performance of each participant was evaluated by calculating the difference between the participant's result and the assigned value ($D_i = x_i - x_{pt}$) [5]. D_i can be interpreted as the measurement error for the results compared to the assigned value. The difference can also be expressed as percentage (D_i%):

$D_i\% = \frac{100\,(x_i - x_{pt})}{x_{pt}}\%$, where x_i = participant's result and x_{pt} = assigned value.

3 Results and conclusions

3.1 Results

The summary of the results of the interlaboratory comparison is shown in Table 5. The results of each participant are given in Appendix 2, the results are presented graphically with their uncertainties in Appendix 3, a summary of the z score are given in Appendix 4, the summary of the answers to the questionnaire is in Appendix 5 and the results grouped according to the methods are presented in Appendix 6.

As expected, the variation of the results in the clear sample was relatively small and all participants, except one, reported satisfactory results.

The variation of the results in the colored sample was larger and unfortunately the issues with homogeneity and stability testing prevented the comparison of the results via z scores.

The D_i % values varied between -8 and 137 (Table 6).

Table 5. The summary of the results in the interlaboratory comparison BTOX 11/2017.

Measurand	Sample	Unit	Assigned value	Mean	Rob. mean	Median	s_{rob}	s_{rob} %	2 x s_{pt} %	n (all)	Acc z %
EC50 value, 15 min	Clear	%	17.3	17.3		17.0			-	4	-
	Color	%		43.1		12.1			-	4	-
EC50 value, 30 min	Clear	%	17.8	17.8	17.4	17.6	2.0	11.5	40	8	88
	Color	%		10.7		10.7			-	6	-

Rob. mean: the robust mean, s_{rob}: the robust standard deviation, s_{rob} %: the robust standard deviation as percent, $2\times s_{pt}$ %: the standard deviation for proficiency assessment at the 95 % confidence level, Acc z %: the results (%), where $|z| \leq 2$, n(all): the number of the participants.

Table 6. The performance of clear sample EC50, 15 min for each participant expressed as D_i %.

Participant	Participant's result (%)	D_i%
2	41	137
3	16	-8
7	19.04	10
8	16.97	-2

3.2 Possible reasons for variance in the colored sample

After extensive development work to find an optimal colored sample for the interlaboratory comparison, we chose the malt beverage based sample for promising results obtained in the preliminary testing. For reasons unknown, the similarly prepared larger sample batch prepared for the interlaboratory comparison round failed to reproduce the same level of homogeneity and stability test results observed in the preliminary tests. However, the sample was relatively stable, as the test results that did not exhibit either raising or lowering trends. Notably large variation in between parallel samples tested within the same day could be result of a pH drift in the sample as the pH value of the sample was checked and adjusted only once per day in the beginning of the testing day and was not rechecked prior to every parallel test.

Another factor contributing to the large variance in the colored sample could be the relatively nontoxic nature of the colored sample. Maximum concentration to be tested in the test procedure was 50 %, because the 100 % native sample had to be diluted to 50 % by the addition of equal volume of bacterial solution. Many of the reported test results were calculated from measurement data with inhibition values below 50 % in the highest concentration of the dilution series. Therefore calculation of EC values was done on the extrapolated area of the inhibition curve. More reliable results are obtained if the EC50 % value is situated between measured dilutions in the dilution series curves. This might further have resulted in larger variations in the accuracy of the predicted EC50 value.

3.3 Results of background questionnaire

Along with the test results, the participants were asked to fill out a questionnaire concerning the details of pretreatments of the samples and testing procedures (Appendix 5). In total, five participants answered to the questionnaire and provided insights to factors that might influence the test results.

The questionnaire revealed differences in the pH adjustment of the samples before measurements. Participant 5 reported a pH value of 6.1 for the clear sample, which is above the recommended threshold for pH adjustment (6.0 – 8.5) according to the standards. In comparison, other participants reported the pH value of said sample to be below 6 and proceeded to adjust the pH value to the middle value of 7.0 ± 0.2. Participant 3 had adjusted the

sample closer to the lower limit at 6.0 ± 0.2. Both methods of adjustments are allowed by the standards, but lower pH may, depending on the sample, result in increased variation in the results between participants.

Both of the results reported by the organizer were produced from samples adjusted to middle value of pH 7.0 ± 0.2. The clear sample results of participants 3 and 5, whose samples were tested at pH 6, did not statistically differ from the results where pH was adjusted to 7.0 ± 0.2 (Appendix 3). However, it is noteworthy that these results performed in lower pH than others, also reported slightly lower results than many other participants, resulting in z scores of -0,51 and -0,28, respectively.

All the participants reported adjusting the pH of the colored sample and again participant 3 adjusted the sample to the lower limit at 6.0 ± 0.2 and others adjusted the sample to the middle value of 7.0 ± 0.2. The results from the colored sample were not evaluated using z scores, but it is evident in the figures in Appendix 3 that participant 3 reported the lowest result. Given the similar observations from the clear sample, we can assume that adjusting the pH value to different pH values is a potential source of variation in comparison test results. When testing natural samples, selection of adjusted pH target closer to real pH value of the sample is more likely to lead to a more realistic test result and toxicity evaluation that includes the possible toxic effect of the pH.

The participants also reported the procedure for adjusting the pH of the 2 % NaCl diluent solution, which is used as a control sample as well as a diluent in preparing the dilution series. In the past, the organizer has had problems with the quality of the control samples if the pH of the diluent has been checked only long beforehand. The organizer observed that the pH of the diluent solution is prone to lower quickly during the testing. Therefore the new procedure is to adjust the diluent pH right prior to use each time a new dilution series is prepared. Some participants reported that the diluent pH was not adjusted at all or only prior to the test. Participant 2 reported adjusting the diluent right before use (Appendix 5). This difference in procedures may affect the level of light production, especially in the low ionic concentration controls and thus affect the test results. In the clear sample the results were similar regardless of pH adjustment or not. Therefore it is likely that the pH adjustment did not have significant on the results.

3.4 Analytical methods

The ecotoxicity test can be performed with either a single tube luminometer or a well plate reader that has luminometer features. The test method itself also offers two options for measurement. The standard method relies on a single measurement result [1] as opposed to the kinetic method, which utilizes continuous measurement for several seconds and a maximum value is used for calculations [2]. The methods also differ on the addition method of the bacteria suspension. Built-in dispenser of the luminometer is used to inject the bacteria suspension to the sample in the kinetic method while it is pipetted in manually in the standard method.

The variations between the results from the clear sample were smaller than between results from the colored sample (Table 4). This was to be expected since the clear sample was a solution of 3,5-DCP in distilled water which has well reproducible response in the test. The colored sample was a mixture of various organic compounds in malt beverage and the homogeneity of the sample was difficult to ensure. In addition to problems related to the pH adjustment of the sample, the dark sample was viscous and could have resulted in errors when preparing the dilution series.

3.5 Uncertainties of the results

Only participants 2 and 5 reported uncertainties for the results. The uncertainty varied between 3-50 % and it was estimated using IQC data from synthetic and routine sample replicates or by other methods not mentioned. The laboratory of SYKE has evaluated the measurement uncertainty for the standard methods by calculating the doubled standard deviation (2×Sd) of the X card for reference substance 3,5-DCP. The 2×Sd uncertainty from test results from years 2016 and 2017 combined was 36 % (n=10). The kinetic method is relatively new and has not collected enough reference sample data yet for reliable uncertainty evaluation. The evaluation of the uncertainty of the methods can be made also using MUkit, the uncertainty calculation program provided by ENVICAL SYKE [6].

4 Evaluation of the results

In regards to the clear sample, the evaluation of participants was based on the z scores, which were interpreted as follows:

Criteria	Performance
$\lvert z \rvert \leq 2$	Satisfactory
$2 < \lvert z \rvert < 3$	Questionable
$\lvert z \rvert \geq 3$	Unsatisfactory

In total, 88 % of the results evaluated based on z scores were satisfactory (Appendix 4) when accepting deviation of 40 % from the assigned value. All (100 %) of the accredited results were satisfactory. In the previous round no z scores were provided [7]. No trends could be detected in D_i % values (Table 6).

The results of the colored sample can only be evaluated as a group instead of comparing z scores in relation to an assigned value. Grouping the results according to test method reveals trends in the order of magnitude of the results (Appendix 6). EC50 30 min results obtained with standard method yielded lower results (1.5 % and 3.2 %) suggesting greater toxicity compared to the kinetic method, where results (10.7 %, 14.1 % and 24.2 %) suggesting lower toxicity. The colored sample was safe for human consumption and as such can be expected to be relatively nontoxic. This trend in EC50 value estimation, although not statistically relevant due to the small dataset (n=5), could be attributed to the better suitability of the kinetic method to

dark and colored samples as the kinetic method allows for several measurements and compensates the loss of light from any light inhibiting color of a sample.

To conclude, this interlaboratory comparison test served as a means for demonstrating the large range of applications applied for the *A. fischeri* - exotoxicity tests in use. In addition, each participant was able to compare the performance of test method to other actors in the field. There is a need for interlaboratory comparison tests with matrices-containing samples and the ecotoxicology laboratory of SYKE will continue to further develop the dark and colored sample and provide these types of interlaboratory comparisons also in the future, as long as there is an interest to take part.

5 Summary

The ecotoxicological laboratory of SYKE and Proftest SYKE carried out this interlaboratory comparison for analysis of ecotoxicity in liquids using *Aliivibrio fischeri* (formerly *Vibrio fischeri*)–bacterial test. One clear synthetic sample and one dark as well as colored sample were sent to participants in October 2017 (BTOX 11/2017). In all, eight participants took part providing 22 results. Both the standard method and the kinetic method were used. Measurements were done in single tube luminometers or in well plate readers that have luminometer features.

The means of the participants' results were used as assigned values for the clear sample. Evaluations of the performances were done by comparing the results of each participant to the assigned values using D_i% values and z scores. As expected, the variation of the results in the clear sample was relatively small and all participants but one reported satisfactory results. In total, 88 % of the results evaluated based on z scores were satisfactory.

The homogeneity of the colored sample was not fully demonstrated. Therefore no evaluations of performances were done. The biggest problems related to credible results were related to pH adjustment of the samples and the non-toxicity of the sample which required extrapolation of the measured results in order to receive an exact value for the EC50%.

6 Summary in Finnish

SYKEn ekotoksikologinen laboratorio ja Proftest SYKE järjestivät akuuttia valobakteeritestiä suorittaville laboratorioille vertailumittauksen lokakuussa 2017 (BTOX 11/2017). Vertailumittaukseen osallistui yhteensä kahdeksan osallistujaa 22 testituloksella. Sekä standardimenetelmää että kineettistä menetelmää käytettiin ja mittaukset suoritettiin putkiluminometrillä tai kuoppalevylukijalla.

Kirkkaan näytteen mittaussuureen vertailuarvona käytettiin osallistujien yksittäisten tulosten keskiarvoa. Osallistujien tuloksia verrattiin vertailuarvoon D_i%- ja z-arvojen avulla. Odotetusti,

osallistujien tulosten hajonnat olivat melko pienet ja kaikkien, paitsi yhden, osallistujan tulokset olivat hyväksyttäviä. Kaikkiaan 88 % tuloksista oli hyväksyttyjä z-arvoja käytettäessä.

Värillisen näytteen homogeenisuutta ei voitu täysin osoittaa ja siksi osallistujien pätevyyksiä ei arvioitu värillisen näytteen kohdalla. Suurimmat tuloksien epävarmuuteen vaikuttavat tekijät olivat näytteen pH:n säätö ja näytteen vähätoksisuus, mikä saattoi kasvattaa osallistuja-kohtaisten tulosten vaihtelua tuloslaskennan ektrapolontivaiheessa.

REFERENCES

1. ISO 11348-3. 2007. Water quality. Determination of the inhibitory effect of water samples on the light emission of *Vibrio fischeri* (Luminescent bacteria test). Part 3: Method using freeze-dried bacteria.
2. ISO 21338. 2010 Kinetic determination of the inhibitory effects of sediment. Other solids and coloured samples on the light emission of Vibrio fischeri (Kinetic luminescent bacteria test).
3. SFS-EN ISO 17043. 2010. Conformity assessment – General requirements for Proficiency Testing.
4. ISO 13528, 2015. Statistical methods for use in proficiency testing by interlaboratory comparisons.
5. Proftest SYKE Guide for laboratories: www.syke.fi/proftest/en Running proficiency test www.syke.fi/download/noname/%7B3FFB2F05-9363-4208-9265-1E2CE936D48C%7D/39886.
6. Näykki, T. Virtanen, A. and Leito, I. 2012. Software support for the Nordtest method of measurement uncertainty evaluation. Accred. Qual. Assur. 17: 603-612. MUkit website: www.syke.fi/envical.
7. Järvistö, J., Meriläinen, P. and Björklöf, K. (2016). Interlaboratory Proficiency Test 14/2015 - Bacterial toxicity test. Reports of the Finnish Environment Institute 30/2016. http://hdl.handle.net/10138/165795.

APPENDIX 1: Terms in the results tables

Results of each participant

Measurand	The tested parameter
Sample	The code of the sample
z score	Calculated as follows: $z = (x_i - x_{pt})/s_{pt}$, where x_i = the result of the individual participant x_{pt} = the assigned value s_{pt} = the standard deviation for proficiency assessment
Assigned value	The value attributed to a particular property of a proficiency test item
2 × s_{pt} %	The standard deviation for proficiency assessment (s_{pt}) at the 95 % confidence level
Participants's result	The result reported by the participant (the mean value of the replicates)
Md	Median
SD	Standard deviation
SD%	Standard deviation, %
n (stat)	Number of results in statistical processing

Summary on the z scores

S – satisfactory ($-2 \leq z \leq 2$)
Q – questionable ($2 < z < 3$), positive error, the result deviates more than 2 × s_{pt} from the assigned value
q – questionable ($-3 < z < -2$), negative error, the result deviates more than 2 × s_{pt} from the assigned value
U – unsatisfactory ($z \geq 3$), positive error, the result deviates more than 3 × s_{pt} from the assigned value
u – unsatisfactory ($z \leq -3$), negative error, the result deviates more than 3 × s_{pt} from the assigned value

Robust analysis

The items of data are sorted into increasing order, $x_1, x_2, x_i, \ldots, x_p$.
Initial values for x^* and s^* are calculated as:
x^* = median of x_i (i = 1, 2, ...,p)
s^* = 1.483 × median of $|x_i - x^*|$ (i = 1, 2, ...,p)

The mean x^* and s^* are updated as follows:
Calculate $\varphi = 1.5 \times s^*$. A new value is then calculated for each result x_i $(i = 1, 2 \ldots p)$:

$$x_i^* = \begin{cases} x^* - \varphi, & \text{if } x_i < x^* - \varphi \\ x^* + \varphi, & \text{if } x_i > x^* + \varphi, \\ x_i & \text{otherwise} \end{cases}$$

The new values of x^* and s^* are calculated from:

$$x^* = \sum x_i^* / p$$

$$s^* = 1.134\sqrt{\sum (x_i^* - x^*)^2 / (p-1)}$$

The robust estimates x^* and s^* can be derived by an iterative calculation, i.e. by updating the values of x^* and s^* several times, until the process convergences [4].

APPENDIX 2: Results of each participant

Participant 1												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 30 min	%	Clear		-0.84	17.8	40	14.8	17.6	17.8	2.8	15.5	7

Participant 2												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 15 min	%	Clear			17.3		41.0	17.0	17.3	1.6	9.0	3
	%	Color					1.0	12.1	43.1	69.9	162.3	4
EC50 value, 30 min	%	Clear		5.67	17.8	40	38.0	17.6	17.8	2.8	15.5	7
	%	Color					1.5	10.7	10.7	9.1	85.2	5

Participant 3												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 15 min	%	Clear			17.3		16.0	17.0	17.3	1.6	9.0	3
	%	Color					3.5	12.1	43.1	69.9	162.3	4
EC50 value, 30 min	%	Clear		-0.51	17.8	40	16.0	17.6	17.8	2.8	15.5	7
	%	Color					3.2	10.7	10.7	9.1	85.2	5

Participant 4												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 30 min	%	Clear		1.60	17.8	40	23.5	17.6	17.8	2.8	15.5	7
	%	Color					24.2	10.7	10.7	9.1	85.2	5

Participant 5												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 30 min	%	Clear		-0.28	17.8	40	16.8	17.6	17.8	2.8	15.5	7

Participant 6												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 30 min	%	Clear		-0.07	17.8	40	17.6	17.6	17.8	2.8	15.5	7
	%	Color					10.7	10.7	10.7	9.1	85.2	5

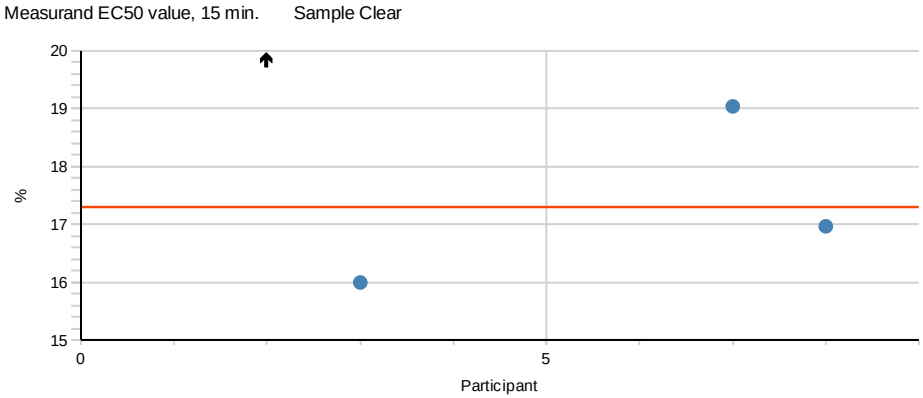
Participant 7												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 15 min	%	Clear			17.3		19.0	17.0	17.3	1.6	9.0	3
	%	Color					20.7	12.1	43.1	69.9	162.3	4
EC50 value, 30 min	%	Clear		0.12	17.8	40	18.2	17.6	17.8	2.8	15.5	7
	%	Color					14.1	10.7	10.7	9.1	85.2	5

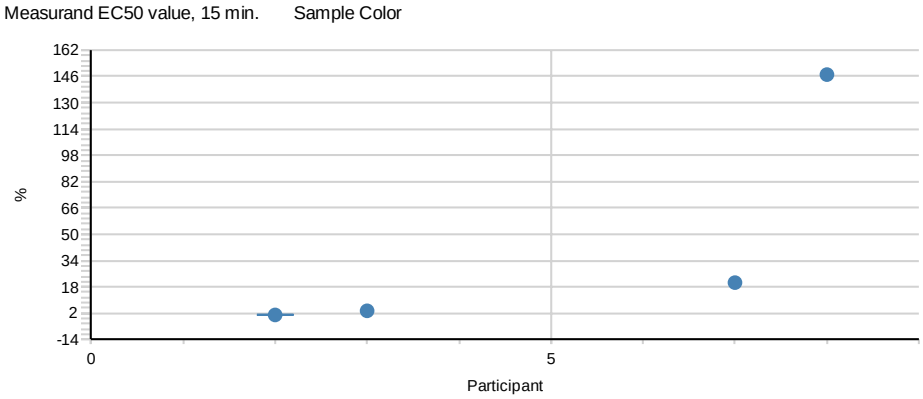
Participant 8												
Measurand	Unit	Sample	-3 0 3	z score	Assigned value	2×s_{pt} %	Participant's result	Md	Mean	SD	SD%	n (stat)
EC50 value, 15 min	%	Clear			17.3		17.0	17.0	17.3	1.6	9.0	3
	%	Color					147.1	12.1	43.1	69.9	162.3	4
EC50 value, 30 min	%	Clear		-0.02	17.8	40	17.7	17.6	17.8	2.8	15.5	7
	%	Color					72.7	10.7	10.7	9.1	85.2	5

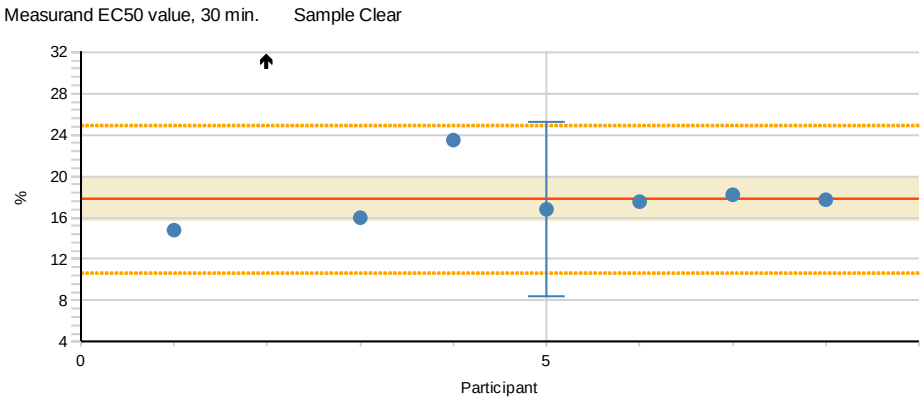
APPENDIX 3: Results of participants and their uncertainties

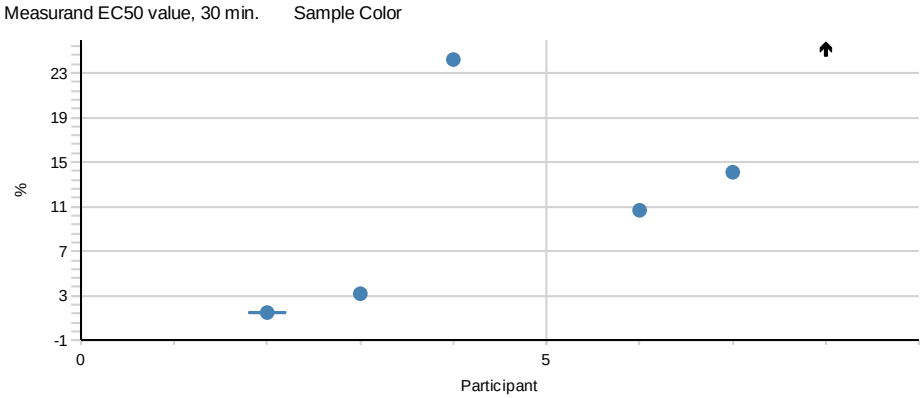
In figures:

- The dashed lines describe the standard deviation for the proficiency assessment, the red solid line shows the assigned value, the shaded area describes the expanded measurement uncertainty of the assigned value, and the arrow describes the value outside the scale.








Measurand EC50 value, 30 min.
Sample Color
23
19
15
11
7
3
-1
%
0
5
Participant

APPENDIX 4: Summary of the z scores

Measurand	Sample	1	2	3	4	5	6	7	8	%
EC50 value, 30 min	Clear	S	*U*	*S*	*S*	S	*S*	*S*	S	87.5
	Color	.	.	.	.	.	.	.	.	
%		100	0	100	100	100	100	100	100	
accredited		1				1			1	

S - satisfactory (-2 ≤ z ≤ 2), Q - questionable (2 < z < 3), q - questionable (-3 < z < -2),
U - unsatisfactory (z ≥ 3), and u - unsatisfactory (z ≤ -3), respectively
bold - accredited, italics - non-accredited, normal - other
% - percentage of satisfactory results

Totally satisfactory, % in all: 88 % in accredited: 100 % in non-accredited: 80

APPENDIX 5: Answers to the questionnaire

A. Basic information on handling of the samples:

1. Original pH of sample

Sample	Original pH
Clear	6.1 5.4 5.42 5.70
Dark	5.4 5.17 4.22

2. Were the pH of the samples adjusted?

a. Clear sample No / Yes (check box), if yes, to what value? (open field)
b. Dark sample No / Yes (check box), if yes, to what value? (open field)

Sample	pH adjusted No	pH adjusted Yes	If yes, to what value?
Clear	20 % (n=1)	80 % (n=4)	6.2 6.8-7.2 7.27 7
Dark	0 % (n=0)	100 % (n=4)	5.9 6.8-7.2 7.53 7

3. Were the samples aerated?

a. Clear sample No / Yes (check box), if yes, to what value? (open field)
b. Dark sample No / Yes (check box), if yes, to what value? (open field)

4. What were the dissolved oxygen contents in the samples?

a. Clear sample (open field)
b. Dark sample (open field)

Sample	Dissolved oxygen contents (mg/l)	Sample aerated No	Sample aerated Yes	If yes , how?
Clear	5.68 10.47	75 % (n=3)	25 % (n=1)	7.07
Dark	2.56	67 % (n=2)	33 % (n=1)	6.20

5. Analysis date

a. Clear sample (open field)
b. Dark sample (open field)

Sample	Analysis Date
Clear	25.10.2017 20.10.2017 19.-20.10.2017 20.10.2017 20.10.2017
Dark	20.10.2017 19.-20.10.2017 20.10.2017 20.10.2017

6. Storing temperature before analysis: (open field)

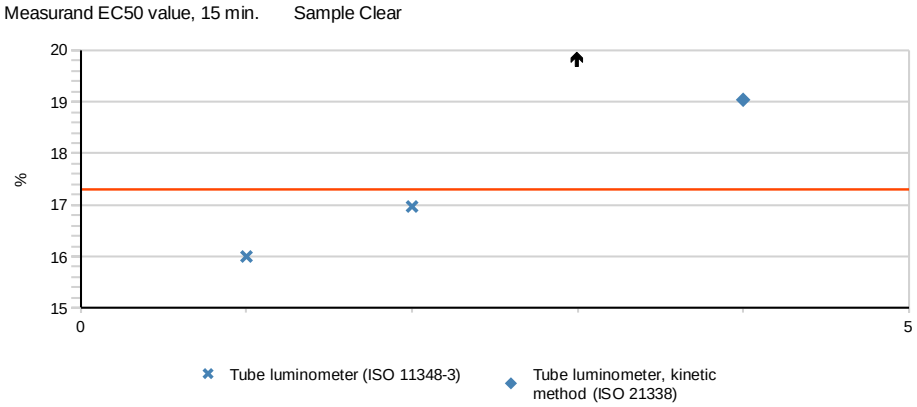
Sample	Storing temp
Clear & Dark	+4 C -20 °C +5 Celsius +4 C 4 C

B. Method information:

Question	Answer 1	Answer 2	Answer 3	Answer 4	Answer 5
If several methods are available, how is the method chosen for any given sample?	We only have this one method available, in use.	Vibrio fischeri SFS-ISO 21338 freeze-dried luminescent bacteria.	We study only kinetic method in our laboratory. We preferred this method because of its reliability for coloured and turbid samples		
If the bacterial suspension is pipetted by hand, what temperature is the pipette calibrated at?	Room temperature.	22 °C	+20 Celsius	We use otomatic pipetting system.	
If the automatic injection method is used for the bacterial suspension applicattion, how is the injector primed?	Not used	200 µL			
Is the pH of the 2% NaCl solution preadjusted or adjusted right before use?	No preadjustion	Preadjusted	No	We don't use NaCl solution, but add always to the samples 0.2 g NaCl grains / 10ml	pH is adjusted to 6,8-7,2 right before use
How many replicates are used in calculating the results?	2	1	Clear: 3 replicates Color: 3 replicates	two replicates	3
What software is used to calculate the results?	Biotox programme .	excel	We don´t use software. Results are calculated by hand in double logarithmic paper	FB12 Sirius Software V2.0	Ascent software + Excel, jossa kaavalaskenta standardin mukaan
How is the uncertainty of the method calculated or assessed?	Using the dicromate and CuSO4 controls and results from parallel samples.	The uncertainty was evaluated from random error caused by parallel samples, because control sample was not available. The random error multiply by 2	Method is not accredited.		

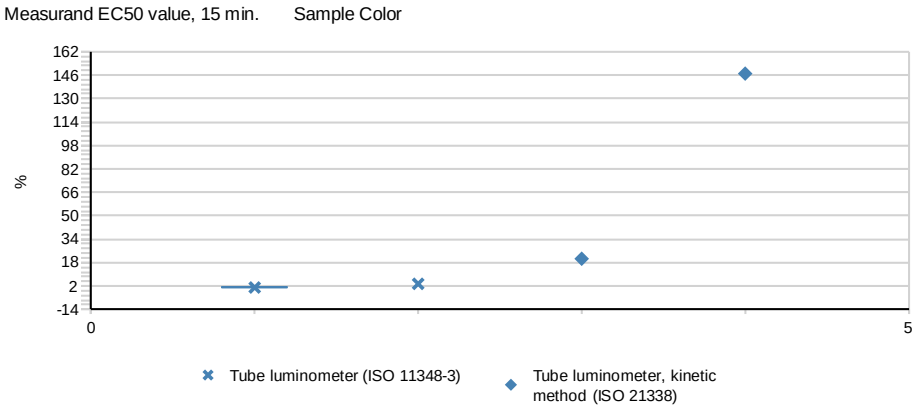
APPENDIX 6: Results grouped according to the methods

The explanations for the figures are described in the Appendix 9. The results are shown in ascending order.



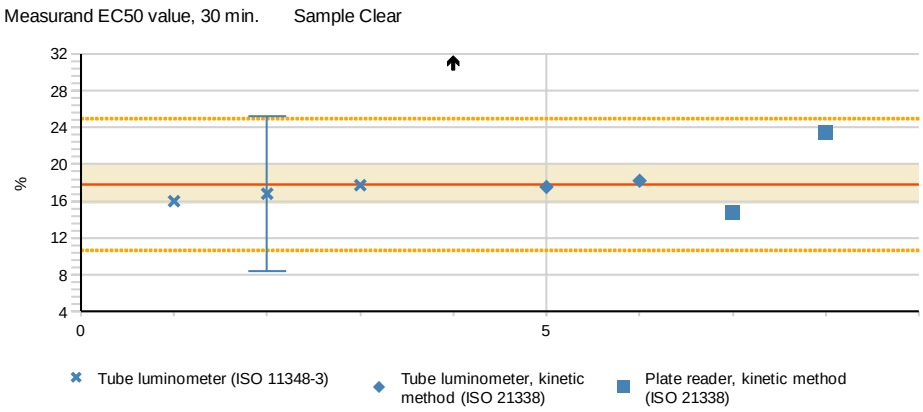
Method	N	Mean (%)	Median (%)	Sd (%)	Min (%)	Max (%)
Tube test, standard method	2	16	16	0.7	16	16.97
Tube test, kinetic method	1	19	19	0	19.04	19.04

N = number of results; Sd = standard deviation; Min =minimum; Max = maximum.



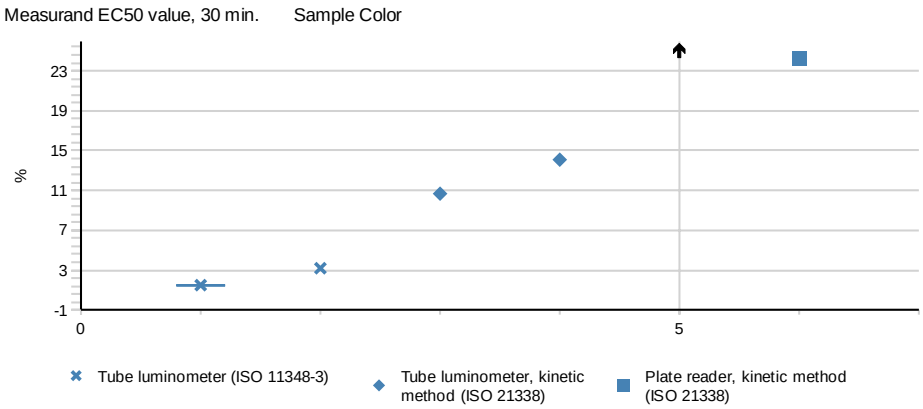
Method	N	Mean (%)	Median (%)	Sd (%)	Min (%)	Max (%)
Tube test, standard method	2	2	2	2	1	3.5
Tube test, kinetic method	2	84	84	89	20.6	147.1

N = number of results; Sd = standard deviation; Min =minimum; Max = maximum.



Method	N	Mean (%)	Median (%)	Sd (%)	Min (%)	Max (%)
Tube test, standard method	3	17	17	1	16	17.8
Tube test, kinetic method	2	18	18	0.5	17.6	18.2
Plate reader, kinetic method	2	19	19	6	14.8	23.5

N = number of results; Sd = standard deviation; Min =minimum; Max = maximum.



Method	N	Mean (%)	Median (%)	Sd (%)	Min (%)	Max (%)
Tube test, standard method	2	2.35	2.35	1.2	1.5	3.2
Tube test, kinetic method	2	12	12	2	10.7	14.1
Plate reader, kinetic method	1	24	24	0	24.2	24.2

N = number of results; Sd = standard deviation; Min =minimum; Max = maximum.

INTERLABORATORY COMPARISON 11/2017

FINNISH ENVIRONMENT INSTITUTE

ISBN 978-952-11-4903-0 (pbk.)
ISBN 978-952-11-4904-7 (PDF)
ISSN 1796-1718 (print)
ISSN 1796-1726 (Online)